

# How does cell size regulation affect population growth?

Jie Lin<sup>1</sup> and Ariel Amir<sup>1</sup>

<sup>1</sup>*School of Engineering and Applied Sciences, Harvard University, Cambridge, Massachusetts 02138, USA*

(Dated: November 28, 2016)

The proliferation of a growing microbial colony is well characterized by the population growth rate. However, at the single-cell level, isogenic cells often exhibit different cell-cycle durations. For evolutionary dynamics, it is thus important to establish the connection between the population growth rate and the heterogeneous single-cell generation time. Existing theories often make the assumption that the generation times of mother and daughter cells are independent. However, it has been shown that to maintain a bounded cell size distribution, cells that grow exponentially at the single-cell level need to adopt cell size regulation, leading to a negative correlation of mother-daughter generation time. In this work, we construct a general framework to describe the population growth in the presence of size regulation. We derive a formula for the population growth rate, which only depends on the variability of single-cell growth rate, independent of other sources of noises. Our work shows that a population can enhance the population growth by reducing the growth rate variability, which may be the reason for the mild growth rate variability observed in many microorganisms.

For exponentially growing populations of microbes, the population growth rate  $\Lambda_p$  characterizes the fitness of the system, measuring how fast the number of individuals  $N(t)$  increases with time  $t$ ,  $N(t) \sim \exp(\Lambda_p t)$ . At the single-cell level, isogenic cells also show diverse phenotypes[1–7], *e.g.*, the generation time for a single cell to divide is heterogeneous in the population[8–12]. The mathematical relation between the heterogeneity of generation time at the single-cell level and the population-level growth rate was investigated initially by Powell[13] with the assumption that each cell has a fluctuating generation time independent of its mother. Powell’s theory does not consider the possibility that cells in consecutive generations can be correlated, hence we refer it as the independent generation time (IGT) model. Because the generation time sets the timing of cell division, the IGT model is also called the timer model[14]. However, recent single-cell level experiments[15–19] demonstrate that the cell volume of many bacteria in fact grows exponentially in time, *e.g.*, *Escherichia coli*, *Caulobacter crescentus*, *Bacillus subtilis*. Exponentially growing cell volume is incompatible with the assumption of the absence of correlations between mother and daughter cells. In the latter case, the random noise in the generation time will lead to divergent fluctuations in the cell volume[20].

Using a discrete Langevin approach, we show that the IGT model is broken down in real systems where the mother-daughter generation times are correlated. Our central result is that the population growth rate is reduced due to the negative mother-daughter correlation compared with the IGT model. Importantly, in the presence of cell size regulation the population growth rate is only determined by the variability of single-cell growth rates, and is independent of the strength of size control or other sources of stochasticity in the generation time. We provide an analytic expression for the population growth rate, which we corroborate by numerical

simulations. Furthermore, we reveal the intricate differences and connections between generation time distributions measured over lineages and population trees, which are experimentally testable.

*The size regulation model*— To make sure cell size is regulated, cells must adopt some regulation strategy[21, 22]. We assume cells attempt to divide symmetrically at the division size,  $v_d = f(v_b)$ . For the sizer model where there is a critical size for cells to divide[23, 24],  $f(v_b) = \text{const}$ , while the timer model corresponds  $f(v_b) = 2v_b$  since the single-cell growth is exponential. Recent experiments[12, 17, 25–27] observe that cell divisions of many microorganisms are in fact regulated by the adder model[20, 27, 28], where cells always attempt to add a constant volume  $v_0$  to its birth size,  $f(v_b) = v_b + v_0$ . It has been argued that the adder model is intimately related to the regulation of genome replication[20, 29].

In the following, we choose a simple regulation model which can unify the three strategies as

$$v_d = 2\alpha v_0 + 2(1 - \alpha)v_b, \quad (1)$$

where  $v_0$  is a constant and  $\alpha$  is the regulation parameter. It is possible to show that  $\alpha = 0, \frac{1}{2}, 1$  correspond respectively to the timer, adder, and sizer model[20]. Given the division size and exponential growth at the single-cell level, we can find the corresponding generation time, for which we consider two sources of noises, the growth rate fluctuations and the time-additive noise  $\xi$ , such that

$$\tau = \frac{1}{\lambda} \ln\left(\frac{v_d}{v_b}\right) + \xi, \quad (2)$$

where the growth rate  $\lambda$  is assumed to be distributed normally with mean  $\lambda_0$  and variance  $\sigma_\lambda^2$ , consistent with various experiments[11, 12, 15, 21]. Similarly, the time-additive noise  $\xi$  is assumed to satisfy a normal distribution with the zero mean and variance  $\sigma_\xi^2$ . Here we assume that during every cell cycle, the growth rate is

a fixed value. Moreover, we assume that the growth rate is independent of growth rates of previous generations, based on the fact that the autocorrelation function of growth rates shows a correlation time less than one generation[15]. We leave the discussion on the effects of correlated growth rate between generations to the Supplemental Material(SM). We point out that the growth rate is not simply  $\ln(2)/(\text{mean generation time})$  as approximated in many cases. In principle, one may also consider size-additive noise; in the SM we show that this does not affect the main conclusion of the present work. When cell size is regulated, the correlation coefficient between the mother-daughter generation time becomes nonzero[12], in contrast to the central assumption of the IGT model. We are able to calculate the correlation of mother-daughter generation time in the limit  $\sigma_\lambda = 0$ ,

$$C_{md} = \frac{\langle \tau_m \tau_d \rangle - \langle \tau_m \rangle \langle \tau_d \rangle}{\sigma_\tau^2} \approx -\frac{\alpha}{2}, \quad (3)$$

where the derivation details are in the SM. For the timer case ( $\alpha = 0$ ), the mother-daughter correlation is zero and the IGT model is valid as expected. However, for  $\alpha > 0$ , the mother-daughter generation time is negatively correlated. A finite  $\sigma_\lambda$  tends to suppress this negative correlation. In the opposite limit  $\sigma_\xi = 0$ ,  $\sigma_\lambda > 0$ , the cell size converges to  $v_0$  immediately and the generation time is merely determined by the growth rate for any  $\alpha$ , in which case the mother-daughter correlation becomes zero and the IGT model is recovered as well.

*A recursive derivation of the independent generation time model.*— In the IGT model, each cell has a random generation time drawn from a given probability distribution  $f(\tau)$ . Imagine the population starts from a single cell at time  $t = 0$ , and after some transient stage, the number of cells increases as  $N(t) \sim \exp(\Lambda_p t)$ . One can alternatively consider the population as two independent populations initiated by the first cell's two daughters. Assuming the first cell divides at time  $\tau$ , then the two sub-populations will increase with time as  $N_1(t) \sim \exp(\Lambda_p(t - \tau))$ . The two interpretations of the population must be equivalent after we average over all possible generation times, so  $\exp(\Lambda_p t) = 2 \int_0^\infty f(\tau) \exp(\Lambda_p(t - \tau)) d\tau$ , leading to the formula[13, 30]

$$2 \int_0^\infty f(\tau) e^{-\Lambda_p \tau} d\tau = 1. \quad (4)$$

It is important to note that we have assumed that the number of cells in the subpopulations increases in the same manner as the original population, except a shift of the beginning time by the mother's generation time.

*A simple case:  $\sigma_\lambda = 0$ .*— Consider a simpler situation without growth rate fluctuations, so all cells grow at the same rate  $\lambda_0$ . One can track a single lineage for many generations, shown as the blue lines in the lineage tree in

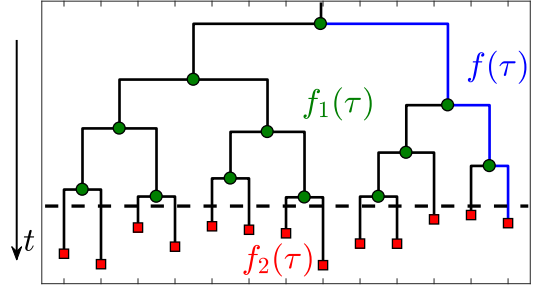


FIG. 1. A lineage tree starting from one cell. Division events are labeled as green circles. To define the generation time distribution, one can track along a single lineage as the blue line, denoted as  $f(\tau)$ . We may also take a snapshot of the population, shown as the dashed line and record all the current and past cells on the tree. Those cells which have already divided live on the branch nodes, so we denote them as branch cells (green circles), and the corresponding generation time distribution is  $f_1(\tau)$ . Those cells which have not divided yet live on the leaf nodes, so we denote them as leaf cells (red squares), and their generation time distribution is called  $f_2(\tau)$ .

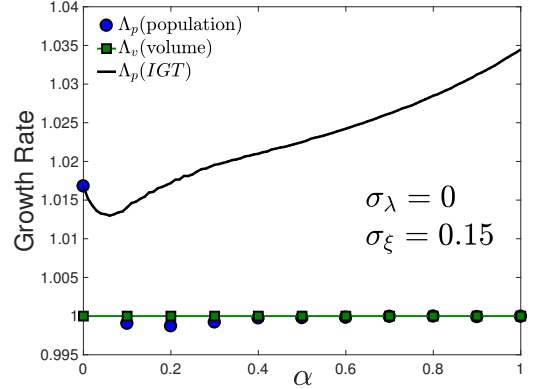


FIG. 2. The IGT model breaks down under cell size regulation ( $\alpha > 0$ ). Numerical simulations of the regulation model, with  $\sigma_\xi = 0.15$ ,  $\sigma_\lambda = 0$ , and  $\lambda_0 = 1$ . The initial number of cells  $N_0 = 1000$ . The relative age of initial cells, the time after the cell's birth divided by the generation time, is uniformly distributed between 0 and 1. Data is collected at time  $t = 10$ . The same numerical protocol is used in the following numerical simulations. The population growth rate is independent of  $\alpha$  for  $\alpha > 0$  with a error bar smaller than the marker size. The systematic small deviations, especially for weak size regulation (small  $\alpha$ ), is presumably due to the finite simulation time. The volume growth rate is always exactly equal to the single cell growth rate  $\lambda_0 = 1$ . The black solid line is the IGT model's prediction, Eq. (4).

Fig. 1. As the generation number increases, the distributions of  $\tau$  and  $v_b$  converge to the stationary distribution. In this case, the average birth volume turns out to be  $\langle v \rangle \approx v_0 \exp(\sigma_v^2/2)$ , with the variance of the birth size distribution equal to  $\sigma_v^2 \approx \lambda_0^2 \sigma_\xi^2 / (2\alpha - \alpha^2)$ [20]. Thus, cell size is regulated to be finite as long as  $0 < \alpha < 2$ . It has been argued that the generation time distribution

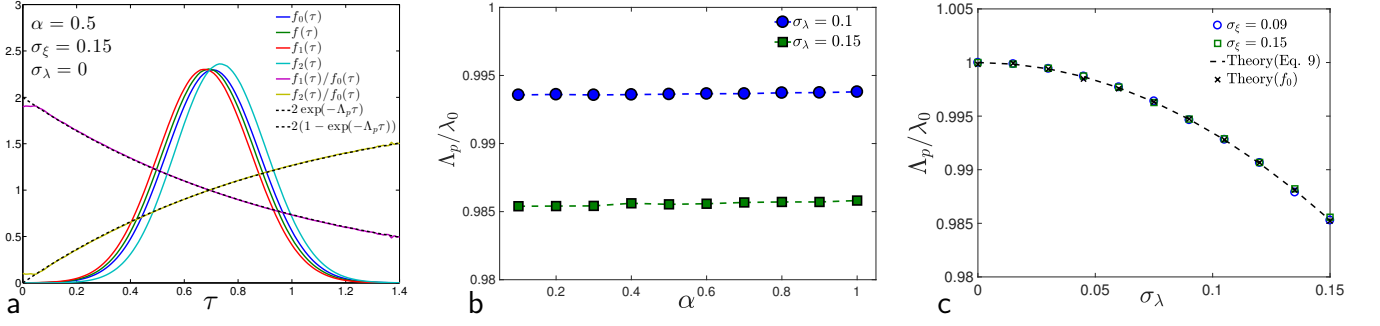


FIG. 3. (a) Generation time distributions, respectively defined for (i) all cells on the tree  $f_0$ , (ii) cells along lineages  $f$ , (iii) only branch cells  $f_1$ , (iv) only leaf cells  $f_2$ , see Fig. 1. Theoretical predictions of the connection between  $f_0$ ,  $f_1$ , and  $f_2$ , Eq. (7),(8) are shown as the dashed lines. (b)  $\Lambda_p/\lambda_0$  v.s.  $\alpha$ , for  $\sigma_\xi = 0.15$ ,  $\lambda_0 = 1$ , and  $\sigma_\lambda$  as indicated. We do not see systematic dependence of  $\Lambda_p$  on  $\alpha$ , and the numerical difference between  $\Lambda_p$  at different  $\alpha$  is of the order  $10^{-4}$ . (c)  $\Lambda_p/\lambda_0$  v.s.  $\sigma_\lambda$  for  $\alpha = 1/2$ ,  $\lambda_0 = 1$ , and  $\sigma_\xi > 0$ . Data with different  $\sigma_\xi$  collapse on the same curve, which means that  $\Lambda_p$  is independent of  $\sigma_\xi$ . The theoretical prediction Eq. (4) matches the numerical results very well. We also show the theoretical prediction from the tree distribution  $f_0$  according to the modified IGT formula,  $2 \int_0^\infty f_0(\tau) e^{-\Lambda_p \tau} d\tau = 1$ . For both (b)&(c), the error bars are smaller than the marker sizes.

along lineages  $f(\tau)$  can be used to infer the population growth rate using the IGT model[30]. As we now show, this is incorrect when there is size regulation.

We first consider the growth of the total cell volume, with a volume growth rate  $\Lambda_v$ ,  $V(t) = \sum_{i=1}^{N(t)} v_i(t) \sim \exp(\Lambda_v t)$ . In the case without single-cell growth rate fluctuations,

$$\frac{dV(t)}{dt} = \sum_{i=1}^{N(t)} \frac{dv_i(t)}{dt} = \lambda_0 V, \quad (5)$$

so  $V(t) = V_0 \exp(\lambda_0 t)$ , where  $V_0$  is the total volume at time  $t = 0$ . The average cell size of the instantaneous population at time  $t$  thus changes as  $\bar{v} \sim \exp((\lambda_0 - \Lambda_p)t)$ . Since the cell size is regulated, one can immediately obtain that  $\Lambda_p = \lambda_0$ , independent of the time-additive noise  $\sigma_\xi$ , and the regulation parameter  $\alpha$ . On the other hand, since the generation time distribution along lineages  $f(\tau)$  turns out to be a normal distribution with mean  $\tau_0 = \ln(2)/\lambda_0$ , and variance  $\sigma_\tau^2 = 2\sigma_\xi^2/(2 - \alpha)$ [20], we can calculate  $\Lambda_p$  as predicted by the IGT model using Eq. (4) for the weak noise limit  $\lambda_0 \sigma_\xi \ll 1$ ,

$$\Lambda_p(IGT) = \frac{\tau_0 - \sqrt{\tau_0^2 - 2 \ln 2 \sigma_\tau^2}}{\sigma_\tau^2}. \quad (6)$$

It is easy to show that  $\Lambda_p(IGT) > \lambda_0$  if  $\sigma_\tau > 0$ . We simulate the model, starting from  $N_0 = 1000$  cells, and compute the population and volume growth rate. Indeed, the volume growth rate  $\Lambda_v$  is exactly equal to  $\lambda_0$  always, and the population growth rate  $\Lambda_p$  matches the volume growth rate for  $\alpha > 0$ , shown in Fig. 2. The IGT model's prediction exceeds the numerical results, except for  $\alpha = 0$  when it is exact. Because  $V(t)$  increases continuously, it is more accurate to quantify in comparison with  $N(t)$ , in

the following numerical results we use the volume growth rate to infer the population growth rate for  $\alpha > 0$ .

*On the generation time distribution.* — Tracking a single lineage is not the only way to define the generation time distribution. Given a lineage tree, one can take a “snapshot” at any time, marked as the dashed line in Fig. 1. To be precise, we need to differentiate two kinds of cells on the tree: (i) cells in the present snapshot (ii) cells that have divided prior to the present snapshot. Since they respectively live on the leaves and branches of the tree, we describe them as leaf cells and branch cells respectively, marked as red squares and green circles in Fig. 1. Besides the distribution along single lineages ( $f(\tau)$ ), we define another three generation time distributions, respectively for the branch cells ( $f_1(\tau)$ ), the leaf cells ( $f_2(\tau)$ ), and all cells on the tree ( $f_0(\tau)$ ). It is only when there is no mother-daughter correlation that  $f(\tau) = f_0(\tau)$ , otherwise, these four distributions are all distinct as shown in Fig. 3(a). Compared with  $f_0(\tau)$ , the leaf cell distribution  $f_2(\tau)$  is biased towards cells with longer generation times because they have a larger chance to be observed in a snapshot. Conversely,  $f_1(\tau)$  is biased towards those cells with shorter generation time[13, 30, 31]. Mathematically,  $f_1(\tau)$ ,  $f_2(\tau)$  are related to  $f_0(\tau)$  by

$$f_1(\tau) = 2e^{-\Lambda_p \tau} f_0(\tau), \quad (7)$$

$$f_2(\tau) = 2(1 - e^{-\Lambda_p \tau}) f_0(\tau). \quad (8)$$

We provide a detailed derivation in the SM. Eqs. (7),(8) are verified for the adder model in Fig. 3(a). Moreover, it turns out that even in the case where there is a finite mother-daughter correlation, one can still make use of the IGT model's result, and the only difference is to replace  $f(\tau)$  by  $f_0(\tau)$  in Eq. (4), as we show in the SM. We test this prediction in the next section.

$\sigma_\lambda > 0$ . — We now turn to the general case with a finite

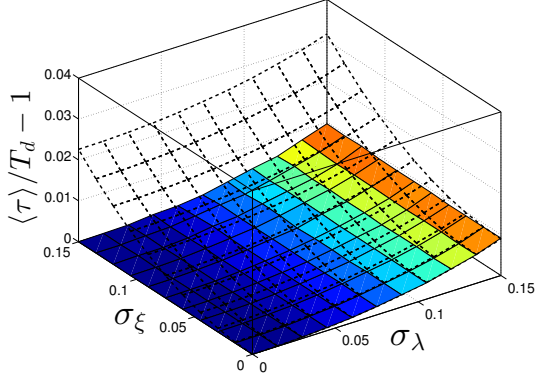


FIG. 4. The relative reduction of population doubling time  $T_d$  against the mean generation  $\langle\tau\rangle$ . The colorful surface plot is the result from the regulation model with  $\alpha = 1/2$ . The dashed surface is the IGT model's prediction using Eq. (4). They coincide only if  $\sigma_\xi = 0$  since the mother-daughter correlation is zero in this case. Otherwise, the correction predicted by the IGT model is much larger than the regulation model's prediction.

variability in the single-cell growth rates. In general, the population growth rate is a function of the three variables  $\Lambda_p(\alpha, \sigma_\xi, \sigma_\lambda)$ , and we have shown that when  $\sigma_\lambda = 0$ ,  $\Lambda_p(\alpha, \sigma_\xi, 0) = \lambda_0$  for any  $\sigma_\xi$ , and  $\alpha > 0$ . Numerically, we find that  $\Lambda_p$  is again independent of  $\alpha$  for a finite  $\sigma_\lambda$ , as shown in Fig. 3(b). In the same way, we find that  $\Lambda_p$  is independent of  $\sigma_\xi$  as well, shown in Fig. 3(c). So  $\Lambda_p$  appears to be a function of  $\sigma_\lambda$ , irrespective of  $\sigma_\xi$  and  $\alpha$ . Given the fact that  $\Lambda_p$  is indeed independent of  $\sigma_\xi$ , we are able to obtain the general expression of  $\Lambda_p$  for real populations since  $\Lambda_p(\alpha, \sigma_\xi, \sigma_\lambda) = \Lambda_p(\alpha, 0, \sigma_\lambda)$ . As we have discussed before, the mother-daughter correlation vanishes in the limit  $\sigma_\xi \rightarrow 0$ , and the generation time is simply  $\tau = \ln(2)/\lambda$ , where  $\lambda$  is assumed to be drawn from a normal distribution. We can calculate the theoretical value of  $\Lambda_p$  using the IGT model's formula Eq. (4) for small  $\sigma_\lambda$  with the saddle point approximation

$$\Lambda_p(\sigma_\lambda) = \lambda_0 \left\{ 1 - \left( 1 - \frac{\ln 2}{2} \right) \left( \frac{\sigma_\lambda}{\lambda_0} \right)^2 \right\}, \quad (9)$$

verified in Fig. 3(c). The detailed derivation is shown in the SM. Eq. (9) implies that one can infer the population fitness indirectly through the growth rate distribution along lineages, without tracking all the cells in the population. Finite growth rate fluctuations tend to *decrease* the population growth rate given a fixed mean value. Experimentally the coefficient of variation (CV) of single-cell growth rate,  $\sigma_\lambda/\lambda_0 \approx 7\%$  has been reported for *E.coli* in different growth conditions[12, 19], much smaller the CV of generation times (20% – 40%)[12, 30]. Our results can explain the origin of the small growth rate variability as it improves the overall fitness of the population. A recent study on the particular case of the sizer model leads to similar conclusions[32]. Surprisingly,

we show here that this extends to the biologically relevant case of the adder model, and in fact, that the results are independent of the strength of size control. The IGT model predicts that the population doubling time  $T_d = \ln 2/\Lambda_p$  is smaller than the mean generation time  $\langle\tau\rangle$  along lineages[30]. In the case where  $\alpha > 0$ , the above conclusion is still valid, shown in the colored surface in Fig. 4. However, the relative difference between the two variables is much more moderate than the IGT model's predictions due to cell size regulation.

*Discussion.*—In this work, we investigated the effects of cell size regulation on population growth. We extend the independent generation time (IGT) model to the more realistic scenario where cell size is controlled. We consider two sources of noises at the single-cell level: (i) growth rate fluctuations, (ii) time-additive noise. We are able to quantify the negative mother-daughter correlation and invalidate the assumption of the IGT model. As long as cells regulate their sizes, the population growth rate has to be equal to the total volume growth rate. We use this identity to prove that the population growth rate is simply equal to the single cell's growth rate in the case without growth rate fluctuations, contradicting the IGT model's predictions. This limit appears to be approximately realized in some biological scenarios[12, 19]. With finite growth rate fluctuations, we find that the population growth rate is again independent of the regulation parameter  $\alpha$ , the strength of time-additive noise, and also size-additive noise as we show in the SM.  $\Lambda_p$  thus only depends on the single-cell growth rate fluctuations, enabling us to calculate the general expression of  $\Lambda_p$ . Furthermore, we clarify the connection between the generation time distributions defined on lineages and trees. Recent microfluidic experiments report a test of IGT model and the authors show that the population doubling time is reduced by a fraction consistent with the IGT model's prediction. However, the experiments are limited to the number of cells in the colonies, which is around 10 cells[30], which make it hard to conclude how much the population growth rate is affected by the mother-daughter correlation. Our work supports that the population doubling time is indeed smaller than the mean generation time, but with a much smaller difference than the IGT model's prediction. In terms of evolution, our model suggests that microbial populations tend to *reduce* the growth rate noise given a constant mean growth rate, which may provide an explanation as to why the growth rate variability is small in certain cases[12, 19].

We thank Sven van Teeffelen, Po-Yi Ho, Felix Wong, and Felix Barber for insightful discussions. AA thanks the A.P. Sloan foundation for their support.

- 
- [1] M. Thattai and A. Van Oudenaarden, Proceedings of the National Academy of Sciences **98**, 8614 (2001).
- [2] S. Di Talia, J. M. Skotheim, J. M. Bean, E. D. Siggia, and F. R. Cross, Nature **448**, 947 (2007).
- [3] A. Raj and A. van Oudenaarden, Cell **135**, 216 (2008).
- [4] H. Salman, N. Brenner, C.-k. Tung, N. Elyahu, E. Stolovicki, L. Moore, A. Libchaber, and E. Braun, Phys. Rev. Lett. **108**, 238105 (2012).
- [5] D. J. Kiviet, P. Nghe, N. Walker, S. Boulineau, V. Sunderlikova, and S. J. Tans, Nature **514**, 376 (2014).
- [6] N. Brenner, E. Braun, A. Yoney, L. Susman, J. Rotella, and H. Salman, The European Physical Journal E **38**, 1 (2015).
- [7] M. Ackermann and F. Schreiber, Environmental microbiology **17**, 2193 (2015).
- [8] O. Sandler, S. P. Mizrahi, N. Weiss, O. Agam, I. Simon, and N. Q. Balaban, Nature **519**, 468 (2015).
- [9] R. Pugatch, Proceedings of the National Academy of Sciences **112**, 2611 (2015).
- [10] N. Rochman, F. Si, and S. X. Sun, Integr. Biol. , (2016).
- [11] M. Wallden, D. Fange, E. G. Lundius, Ö. Baltekin, and J. Elf, Cell **166**, 729 (2016).
- [12] S. Taheri-Araghi, S. Bradde, J. T. Sauls, N. S. Hill, P. A. Levin, J. Paulsson, M. Vergassola, and S. Jun, Current Biology **25**, 385 (2015).
- [13] E. Powell, Microbiology **15**, 492 (1956).
- [14] J. T. Sauls, D. Li, and S. Jun, Current opinion in cell biology **38**, 38 (2016).
- [15] P. Wang, L. Robert, J. Pelletier, W. L. Dang, F. Taddei, A. Wright, and S. Jun, Current Biology **20**, 1099 (2010).
- [16] M. Godin, F. F. Delgado, S. Son, W. H. Grover, A. K. Bryan, A. Tzur, P. Jorgensen, K. Payer, A. D. Grossman, M. W. Kirschner, *et al.*, Nature methods **7**, 387 (2010).
- [17] M. Campos, I. V. Surovtsev, S. Kato, A. Paintdakhi, B. Beltran, S. E. Ebmeier, and C. Jacobs-Wagner, Cell **159**, 1433 (2014).
- [18] S. Iyer-Biswas, C. S. Wright, J. T. Henry, K. Lo, S. Burrov, Y. Lin, G. E. Crooks, S. Crosson, A. R. Dinner, and N. F. Scherer, Proceedings of the National Academy of Sciences **111**, 15912 (2014).
- [19] N. Cermak, S. Olcum, F. F. Delgado, S. C. Wasserman, K. R. Payer, M. A. Murakami, S. M. Knudsen, R. J. Kimmerling, M. M. Stevens, Y. Kikuchi, *et al.*, Nature Biotechnology (2016).
- [20] A. Amir, Phys. Rev. Lett. **112**, 208102 (2014).
- [21] M. Osella, E. Nugent, and M. C. Lagomarsino, Proceedings of the National Academy of Sciences **111**, 3431 (2014).
- [22] L. Robert, M. Hoffmann, N. Krell, S. Aymerich, J. Robert, and M. Doumic, BMC biology **12**, 1 (2014).
- [23] S. Cooper, *Bacterial growth and division: biochemistry and regulation of prokaryotic and eukaryotic division cycles* (Elsevier, 1991).
- [24] A. L. Koch, *Bacterial growth and form* (Springer, Berlin, 2001).
- [25] M. Deforet, D. van Ditmarsch, and J. B. Xavier, Biophysical Journal **109**, 521 (2015).
- [26] I. Soifer, L. Robert, and A. Amir, Current Biology **8**, 356 (2016).
- [27] A. S. Kennard, M. Osella, A. Javer, J. Grilli, P. Nghe, S. J. Tans, P. Cicuta, and M. Cosentino Lagomarsino, Phys. Rev. E **93**, 012408 (2016).
- [28] K. R. Ghusinga, C. A. Vargas-Garcia, and A. Singh, Scientific Reports **6** (2016).
- [29] P.-Y. Ho and A. Amir, Frontiers in microbiology **6** (2015).
- [30] M. Hashimoto, T. Nozoe, H. Nakaoka, R. Okura, S. Akiyoshi, K. Kaneko, E. Kussell, and Y. Wakamoto, Proceedings of the National Academy of Sciences **113**, 3251 (2016).
- [31] Y. Wakamoto, A. Y. Grosberg, and E. Kussell, Evolution **66**, 115 (2012).
- [32] A. Olivier, arXiv:1602.06970 (2016).

## SUPPLEMENTAL MATERIAL

### MOTHER-DAUGHTER CORRELATION OF GENERATION TIME

To make the theoretical analysis feasible, we also consider an approximate version of the regulation model[20],

$$v_d = 2v_0^\alpha v_b^{1-\alpha}, \quad (\text{A.1})$$

and the corresponding generation time becomes

$$\tau = \frac{\ln 2}{\lambda} - \frac{\alpha}{\lambda} \ln\left(\frac{v_b}{v_0}\right) + \xi. \quad (\text{A.2})$$

It is easy to show that the approximate expression of  $v_d$  shares the same slope as the original model at  $v_b = v_0$ . Because the coefficient of variation of cell birth sizes are often reported to be around 10%[17], this model provides a very good approximation of the original regulation model in the main text. As the generation number increases, the distributions of  $\tau$  and  $v_b$  converge to the stationary distribution. The approximate model enables us to find the distribution analytically along single lineages. Both  $f(\tau)$  and  $\rho(\ln(v_b/v_0))$  satisfy normal distributions, with mean  $\ln 2/\lambda_0$ , 0, and variance  $\sigma_\tau = 2\sigma_\xi^2/(2-\alpha)$ ,  $\sigma_v^2 = \lambda_0^2\sigma_\xi^2/(2\alpha-\alpha^2)$ [20], respectively. Once we have regulation,  $\alpha > 0$ , the assumption of no mother-daughter correlation breaks down. Here we calculate the mother-daughter correlation, using the approximate model. We first consider the case without growth rate fluctuations. At generation  $n+1$ ,  $\tau(n+1)$  is determined by  $v_b(n)$  and  $\tau(n)$  as

$$\tau(n+1) = \frac{\ln(2)}{\lambda_0} - \frac{\alpha}{\lambda_0} \ln\left(\frac{v_b(n)e^{\lambda_0\tau(n)}}{2v_0}\right) + \xi_{n+1}. \quad (\text{A.3})$$

The auto-correlation between  $\tau(n)$  and  $\tau(n+1)$  becomes

$$\langle \tau(n+1)\tau(n) \rangle_c = -\frac{\alpha}{\lambda_0} \langle \ln\left(\frac{v_b}{v_0}\right)\tau \rangle_c - \alpha\sigma_\tau^2, \quad (\text{A.4})$$

where  $\langle AB \rangle_c = \langle AB \rangle - \langle A \rangle \langle B \rangle$ . Using Eq. (A.2), we get  $\langle \ln(v_b/v_0)\tau \rangle_c = -\alpha\sigma_v^2/\lambda_0$ , and obtain the correlation coefficient between the mother-daughter generation time as

$$C_{md} = \frac{\langle \tau(n+1)\tau(n) \rangle_c}{\sigma_\tau^2} = -\frac{\alpha}{2}. \quad (\text{A.5})$$

Indeed there is negative correlation between the mother and daughter. More complicated calculations including a small but finite  $\sigma_\lambda$  can be done. Importantly, finite growth rate fluctuations reduce the negative correlation. On the other hand, in the limit  $\sigma_\xi = 0$ ,  $\sigma_\lambda > 0$ , the fluctuations of cell size are completely erased: every cell is born at  $v_0$ , and divides at  $2v_0$ . But the generation time is still stochastic and determined by the growth rate, so the mother-daughter correlation becomes zero. We numerically calculate  $C_{md}$  using the original regulation model, with two parameters  $\sigma_\xi$ ,  $\sigma_\lambda$ , shown in Fig. A.1.

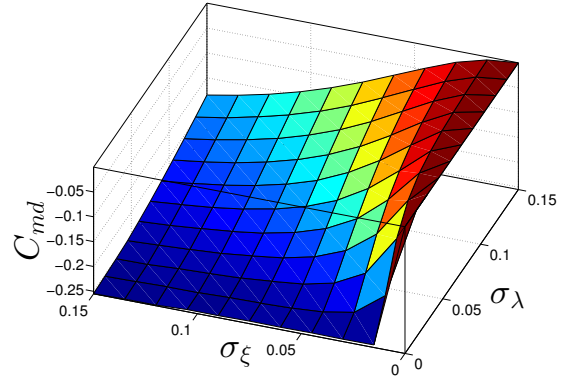


FIG. A.1. Mother-daughter correlation of generation time  $C_{md}$ , as function of the time-additive noise amplitude  $\sigma_\xi$  and the growth rate fluctuation amplitude  $\sigma_\lambda$  for the exact regulation model with regulation parameter  $\alpha = 0.5$ .  $C_{md} \approx -0.25$  for  $\sigma_\xi > 0$ ,  $\sigma_\lambda = 0$ , and  $C_{md} = 0$  for  $\sigma_\xi = 0$ ,  $\sigma_\lambda > 0$ .

### DERIVATION OF THE POPULATION GROWTH RATE

In the limit  $\sigma_\xi \rightarrow 0$ , and a small but finite  $\sigma_\lambda \ll \lambda_0$ , we are able to calculate the analytic expression of the population growth rate  $\Lambda_p$  using the formula of the IGT model. In this case,  $\tau = \ln(2)/\lambda$  because the birth size  $v_b$  converge to  $2v_0$  very quickly without time-additive or size-additive noise, and there are essentially no size fluctuations. We can redefine the variable  $x = \ln(2)/\tau$ , and the formula to calculate  $\Lambda_p$  becomes

$$2 \int_0^\infty \frac{1}{\sqrt{2\pi\sigma_\lambda^2}} \exp\left(-\frac{\ln(2)\Lambda_p}{x}\right) \exp\left(-\frac{(x-\lambda_0)^2}{2\sigma_\lambda^2}\right) dx = 1. \quad (\text{A.6})$$

We can rewrite the integral without the prefactor as

$$\begin{aligned} I &\approx \int_0^\infty \exp\left(-\frac{1}{2\sigma_\lambda^2}((x-\lambda_0)^2 + \frac{2\sigma_\lambda^2\Lambda_p \ln(2)}{x})\right) dx \\ &= \int_0^\infty \exp\left(-\frac{1}{2\sigma_\lambda^2}g(x)\right) dx, \end{aligned} \quad (\text{A.7})$$

and use the saddle point method to calculate the integral,

$$I = \exp\left(-\frac{1}{2\sigma_\lambda^2}g(x_c)\right) \sqrt{\frac{4\pi\sigma_\lambda^2}{g''(x_c)}}. \quad (\text{A.8})$$

$x_c$  is determined by  $g'(x_c) = 0$ , from which we get  $x_c \approx \lambda_0 + \frac{\sigma_\lambda^2\Lambda_p \ln(2)}{\lambda_0^2}$  and

$$g(x_c) \approx \frac{2\sigma_\lambda^2\Lambda_p \ln(2)}{\lambda_0 + \frac{\sigma_\lambda^2\Lambda_p \ln(2)}{\lambda_0^2}} + \frac{\sigma_\lambda^4\Lambda_p^2 \ln(2)^2}{\lambda_0^4}, \quad (\text{A.9})$$

$$g''(x_c) \approx 2 + \frac{4\sigma_\lambda^2\Lambda_p \ln(2)}{(\lambda_0 + \frac{\sigma_\lambda^2\Lambda_p \ln(2)}{\lambda_0^2})^3}. \quad (\text{A.10})$$



Keeping the lowest correction due to  $\sigma_\lambda$ , we eventually find

$$\Lambda_p = \lambda_0 \left(1 - \left(1 - \frac{\ln(2)}{2}\right) \left(\frac{\sigma_\lambda}{\lambda_0}\right)^2\right). \quad (\text{A.11})$$

### EFFECTS OF SIZE-ADDITIVE NOISE

Besides time-additive noise, we also test the effects of size-additive noise in the following way,

$$v_d = 2\alpha v_0 + 2(1 - \alpha)v_b + 2\eta, \quad (\text{A.12})$$

here  $\eta$  is the size-additive noise, with zero mean and variance  $\sigma_\eta^2$ . We fix the time-additive noise as  $\sigma_\xi = 0.05$ , and change  $\sigma_\eta$ ,  $\sigma_\lambda$ . Numerical simulations support that the size-additive noise does not affect the population growth rate, and the numerical values can again be predicted by Eq. (A.11), as shown in Fig. A.2.

### RELATIONS BETWEEN DIFFERENT GENERATION TIME DISTRIBUTIONS

#### Two kinds of cells: branch cell and leaf cell

As shown in Fig. 1 in the main text, there are two different kinds of cells on the lineage tree, which are respectively the current generation (leaf cells) and cells from previous generations (branches cells). The following identity is very useful:  $N_{leaf} = N_{branch} + 1$ , true for any lineage tree with binary division. Besides the generation time distribution along a lineage  $f(\tau)$ , we furthermore define  $f_0(z)$ ,  $f_1(z)$ ,  $f_2(z)$  as the distributions of some cell cycle quantities  $z$ , respectively for all cells in the tree (tree distribution), only branch cells (branch distribution), and only leaf cells (leaf distribution). Because for the whole populations  $N_{leaf} \approx N_{branch}$ , it is always true that

$$f_0(z) = \frac{f_1(z) + f_2(z)}{2}. \quad (\text{A.13})$$

#### Age distribution: $\phi(x)$

In the exponential growth phase of the population, each cell essentially plays the same role and can be considered as the leading cell of the following lineage tree. One can therefore use the tree distribution  $f_0(\tau)$  as the generation time distribution of each cell on the tree, and if there is no mother-daughter correlation,  $f_0(\tau)$  is equal to  $f(\tau)$ . In the following, we essentially follow the same idea originally by Powell[13, 30], and the key difference is to replace the lineage distribution  $f(\tau)$  by the tree distribution  $f_0(\tau)$ .

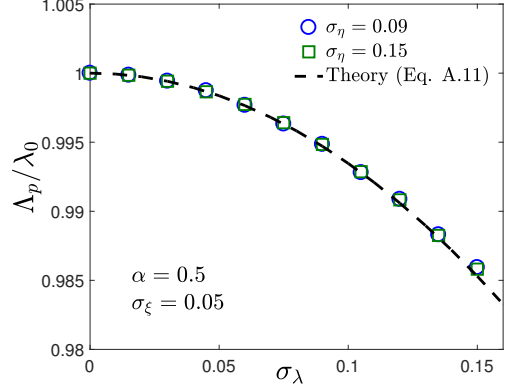


FIG. A.2.  $\Lambda_p$  v.s.  $\sigma_\lambda$ , with two different size-additive noises  $\sigma_\eta$ . Here, we fix the time-additive noise as  $\sigma_\xi = 0.05$ .

Given  $f_0(\tau)$ , we define the survival function, which gives the probability for cells to survive at least to the age  $x$

$$F_-(x) = \int_x^\infty f_0(y)dy, \quad (\text{A.14})$$

and the division rate as

$$\mu(x) = \frac{F_-(x) - F_-(x+dx)}{F_-(x)dx} = \frac{f_0(x)}{F_-(x)}. \quad (\text{A.15})$$

We can alternatively express  $F_-(x)$  as  $F_-(x) = \exp(-\int_0^x \mu(y)dy)$ .

For large time, the age distribution  $\phi(x, t)$  will approach a stationary distribution  $\phi(x)$ . The number of cells at time  $t$  at age  $x$  is thus  $N(t)\phi(x)$ , and the probability for them to live to  $t+dt$  is  $1 - \mu(x)dt$ . In the exponential phase, the number of cells increases exponentially as  $N(t) \sim \exp(\Lambda_p t)$ , with the population growth rate  $\Lambda_p$ . So we can find the self-consistent equation for the age distribution

$$\frac{N(t)\phi(x)(1 - \mu(x)dt)}{N(t+dt)} = \phi(x+dt), \quad (\text{A.16})$$

so we get the equation of  $\phi(x)$  as

$$\frac{d\phi(x)}{dx} = -\mu(x)\phi(x) - \Lambda_p\phi(x). \quad (\text{A.17})$$

Using  $F_-(x) = \exp(-\int_0^x \mu(y)dy)$ , we find the solution as  $\phi(x) = \phi_0 \exp(-\Lambda_p x) F_-(x)$  with  $\phi_0$  to be determined. The population growth rate can be expressed as the integral of age distribution and division rate

$$\begin{aligned} \Lambda_p &= \frac{N(t) \int_0^\infty \phi(x) \mu(x) dx}{N(t)} \\ &= \int_0^\infty \phi(x) \frac{f_0(x)}{F_-(x)} dx = \int_0^\infty \phi_0 e^{-\Lambda_p x} f_0(x) dx, \end{aligned} \quad (\text{A.18})$$

because it is the cells' divisions that increases the number of cells.

From the normalization of  $\phi(x)$ , we get

$$\begin{aligned} 1 &= -\frac{\phi_0}{\Lambda_p} \left\{ -1 + \int_0^\infty e^{-\Lambda_p x} f_0(x) dx \right\} \\ &= \frac{\phi_0}{\Lambda_p} - \frac{\phi_0}{\Lambda_p} \int_0^\infty e^{-\Lambda_p x} f_0(x) dx. \end{aligned} \quad (\text{A.19})$$

Combining with Eq. (A.18), we get  $\phi_0 = 2\Lambda_p$ , and

$$\phi(x) = 2\Lambda_p e^{-\Lambda_p x} F_-(x) \quad (\text{A.20})$$

#### Ancestors' generation time distribution: $f_1(\tau)$

In this subsection, we derive the expression of  $f_1(\tau)$ , given tree distribution  $f_0(x)$ . Imagine we review a snapshot at some time  $t'$  before the current time  $t$ , and all the cells at time  $t'$  are branch cells. Let's first write down the expression of  $f_1$  and explain its interpretation,

$$f_1(\tau) = \frac{N(t')\phi(\tau)\mu(\tau)d\tau}{N(t')\Lambda_p d\tau} \quad (\text{A.21})$$

Here, the denominator represents the number of division events from  $t'$  to  $t' + d\tau$ . The numerator represents cells which divide at age  $\tau$  at time  $t'$ . Their ratio becomes the fraction of branch cells that divide at age  $\tau$ , which means that they have a generation time equal to  $\tau$ . The above derivation does not apply to leaf cells, since they have not divided yet. One should note that here we are not making any assumptions about the correlation between mother-daughter generation time since we focus on one single generation, so Eq. (A.21) is a universal result. Using the expression of  $\phi(x)$ , we obtain

$$f_1(\tau) = 2e^{-\Lambda_p \tau} f_0(\tau), \quad (\text{A.22})$$

and we get the equation to extract the population growth rate  $\Lambda_p$  from the normalization of  $f_1(\tau)$ ,

$$\int_0^\infty 2e^{-\Lambda_p \tau} f_0(\tau) d\tau = 1, \quad (\text{A.23})$$

as verified in the main text.

#### Current cells' generation time distribution: $f_2(\tau)$

We first write down the expression of  $f_2(\tau)$

$$f_2(\tau) = \frac{\int_0^\tau N(t)\phi(y)\frac{F_-(\tau)}{F_-(y)}\mu(\tau)dy}{N(t)}, \quad (\text{A.24})$$

here the nominator means those cells who are at age  $y$  at time  $t$  and will survive until age  $\tau$  and divide. Given  $\phi(\tau)$ , it is easy to obtain

$$f_2(\tau) = 2f_0(\tau)(1 - e^{-\Lambda_p \tau}). \quad (\text{A.25})$$

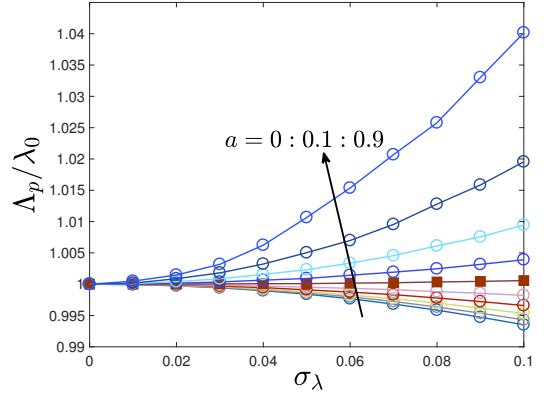


FIG. A.3.  $\Lambda_p$  v.s.  $\sigma_\lambda$ , with fixed time-additive noise  $\sigma_\epsilon = 0.15$ , and zero size-additive noise. The mean value of single-cell growth rate is fixed to be 1, so  $b = 1 - a$ .  $\sigma_\epsilon$  is adjusted according to  $\sigma_\lambda$ . The parameter  $a$  introduced in Eq. (A.27), changes from 0 to 0.9 with step 0.1 as indicated in the figure, and  $a = 0.5$  is shown in filled square.

Again here we do not make any assumption about the mother-daughter correlation. From Eqs. A.22, A.25, we indeed find that

$$f_0(\tau) = \frac{f_1(\tau) + f_2(\tau)}{2}. \quad (\text{A.26})$$

#### EFFECTS OF GROWTH RATE CORRELATIONS

To consider the effect of the correlation of single-cell growth rates between consecutive generations, we assume the daughter cell's growth rate is dependent on its mother's growth rate, and also subject to random noise,

$$\lambda(n+1) = a\lambda(n) + b + \epsilon \quad (\text{A.27})$$

here  $a$  is a number between 0 and 1,  $\epsilon$  is the growth rate noise, with variance  $\sigma_\epsilon^2$ . The mean value of growth rate along lineages becomes  $\langle \lambda \rangle = b/(1 - a)$ , with variance  $\sigma_\lambda^2 = \sigma_\epsilon^2/(1 - a^2)$ . The correlation coefficient between the mother-daughter growth rates is equal to  $a$ . It is clear that if  $a = 0$ , we go back to the case without growth rate correlation, and if  $a \rightarrow 1$ , the variance of growth rate diverges. So far, it is not clear what value  $a$  takes in real microorganisms, and small values of  $a$  have been reported[15]. Analytic computations are hard to do due to the finite correlation of growth rate between generations. Our numerical simulations show that a positive correlation of growth rate tends to increase the population growth rate, as shown in Fig. A.3. When  $a \gtrsim 0.5$ , the population growth rate even increases as the growth rate variability increases. We expect more systematic experiments to quantify the correlation of single-cell growth rates. We cannot exclude the possibility that in some cases, a strong correlation of growth rates may favor a larger growth rate variability.